Sitecore with SOLR

SUGCH – A brief overview of SOLR

Presented by // Julian D. Bagdahn

27 April 2015

Agenda

- What is SOLR?

- Consider SOLR, if...

- What is SolrCloud?

- SOLR in context with Sitecore

- Where to install

- Configuration

- Nice to know

Sitecore Search

Search works on an *index* that is *external to Sitecore*

- Configuration in **/App_Config/Include/** (default and index-specific)
- Determine how items should be indexed

Uses provider model

- Identical *LINQ layer*
- Provider *translates* to work with chosen search technology

Which one should I use?

- Sitecore ships with **Lucene** and **Solr** (**Solr** wraps **Lucene** and is more advanced)
- Use **Solr** to *scale*

What is SOLR

Apache Solr is a standalone enterprise search server with a REST-like API. You put documents in it (called "indexing") via JSON, XML, CSV or binary over HTTP. You query it via HTTP GET and receive JSON, XML, CSV or binary results.

It adds **several features** like:

- Advanced Full-Text Search Capabilities
- Optimized for High Volume Traffic
- Standards Based Open Interfaces - XML, JSON and HTTP
- Comprehensive Administration Interfaces
- Easy Monitoring
- Highly Scalable and Fault Tolerant
- Caching
- Replication

Unlike Lucene, Solr is a web application (WAR) which can be deployed in any servlet container, e.g. Jetty, Tomcat, Resin, etc.

Consider SOLR, if

- You need to index a large number of items, e.g. 50 thousand upwards, Solr has improved query caching and is designed for large numbers of items

- The resilience of the search tier is of maximum business importance, i.e. the site is purely driven by search. Solr provides a more robust replication/sharding and failover system with SolrCloud

- Re-purposing of the search tier in other application is important (non Sitecore) - Solr is a search application and as such can be accessed over HTTP with XML/JSON etc., which makes integration with external systems easier

- You need specific additional Solr features that Lucene doesn't offer

- Your Sitecore environment contains clustered, load-balanced content delivery servers

What is SolrCloud ?

- On a single instance, Solr has something called a SolrCore that is essentially a single index

- If you want multiple indexes, you create multiple SolrCores

- With SolrCloud, a single index can span multiple Solr instances

- This means that a single index can be made up of multiple SolrCore's on different machines

- We call all of these SolrCores that make up one logical index a collection

- A collection is a essentially a single index that spans many SolrCore's, both for index scaling as well as redundancy

If you are already using Lucene on a scaled environment, you <u>MUST</u> use SOLR!

SOLR in context with Sitecore

- SOLR can be used everywhere instead of standard Lucene

- Especially in Sitecore XP 8, where all analytics features as well as e.g. FXM uses an own index, SOLR is the way to go

- SOLR is used via the search provider model in Sitecore, switching from Lucene over to SOLR has no impact on your search code (LINQ) written for Lucene
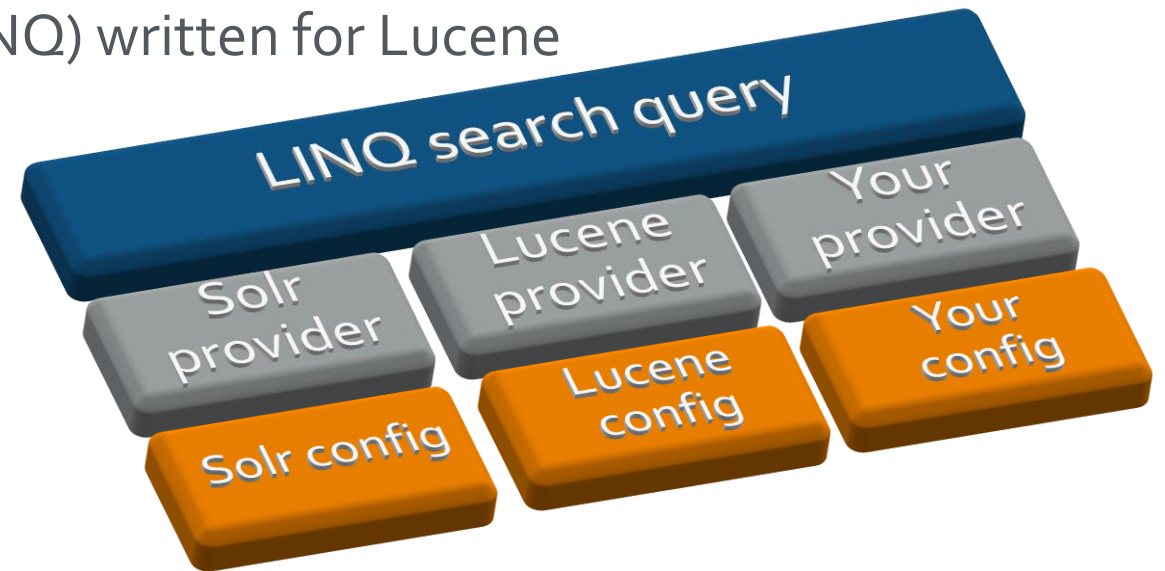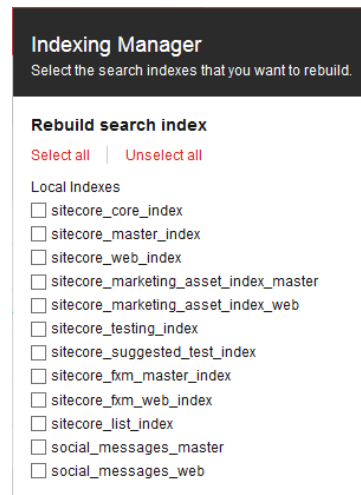
Where to install?

- As SOLR needs an own webserver and servlet container, e.g. Apache Tomcat it is hardly recommended to set up a new, separate machine

- The SOLR instance may receive a lot of requests to process, especially if you have chosen SOLR to use it with some other applications

- Therefore take care of your machines setup, you will find a lot of good examples at the very active SOLR community

- On development environments you are fine to use Bitnami for setting up a local SOLR instance

Configuration

- One SOLR core for each index

- Use Sitecore's SOLR Schema.xml generator to setup the cores schema configuration

- Inside App_Config/Include locate every config-file with „Lucene" or „Solr" in its name

- Remove the *.example* or *.disabled* suffix from each Solr config file and append *.disabled* to each Lucene config file

- Within "Sitecore.ContentSearch.Solr.DefaultIndexConfiguration.config", update the value attribute of the following element with the address of your Solr server:

  ```
  <setting name="ContentSearch.Solr.ServiceBaseAddress" value="http://localhost:8983/solr" />
  ```

- Copy all the .dll files from the **Sitecore Solr Support package** to Sitecore's /bin directory.

- Go to Sitecore's Control Panel and rebuild all indexes using the "Indexing Manager"

Nice to know

- Since SOLR 4.8+ the structure of the default schema.xml has changed, therefore you need to adjust the SOLR schema.xml for each core before sending it through Sitecore's schema.xml generator

  - See https://kb.sitecore.net/articles/227897

- In the **Sitecore community docs** you will find a lot of information regarding SOLR and its setup and configuration, please feel free to contribute:

  - See https://sitecore-community.github.io/docs/search/solr/

- You need to download the **Sitecore SOLR Support Package**, for XP 8 you will find the package in the download section of http://dev.sitecore.net/ under "Additional Tools"

- Questions regarding SOLR in context with Sitecore are covered by our professional support

Thank you for your attention!